Organization TC1600 Bldg./Room REMSEN

U. S. DEPARTMENT OF COMMERCE
PATENT AND TRADEMARK OFFICE
WASHINGTON, DC 20231
IF UNDELIVERABLE RETURN IN TEN DAYS

OFFICIAL BUSINESS

AN EQUAL OPPORTUNITY EMPLOYER

UNITED STATES PATENT AND TRADEMARK OFFICE

UNITED STATES DEPARTMENT OF COMMERCE
United States Patent and Trademark Office
Address: COMMISSIONER FOR PATENTS
P.O. Box 1450
Alexandria, Virginia 22313-1450
www.uspto.gov

APPLICATION NO.	FILING DATE	FIRST NAMED INVENTOR	ATTORNEY DOCKET NO.	CONFIRMATION NO.
09/421,422	10/19/1999	PEHR B. HARBURY	8600-0197.30	4130

22918 7590 05/20/2004

PERKINS COIE LLP
P.O. BOX 2168
MENLO PARK, CA 94026

EXAMINER	
TRAN, MY CHAU T	
ART UNIT	**PAPER NUMBER**
1639	

DATE MAILED: 05/20/2004

Please find below and/or attached an Office communication concerning this application or proceeding.

Office Action Summary	Application No.	Applicant(s)	
	09/421,422	HARBURY ET AL.	
	Examiner	Art Unit	
	MY-CHAU T TRAN	1639	

-- The MAILING DATE of this communication appears on the cover sheet with the correspondence address --

Period for Reply

A SHORTENED STATUTORY PERIOD FOR REPLY IS SET TO EXPIRE 3 MONTH(S) FROM THE MAILING DATE OF THIS COMMUNICATION.

- Extensions of time may be available under the provisions of 37 CFR 1.136(a). In no event, however, may a reply be timely filed after SIX (6) MONTHS from the mailing date of this communication.
- If the period for reply specified above is less than thirty (30) days, a reply within the statutory minimum of thirty (30) days will be considered timely.
- If NO period for reply is specified above, the maximum statutory period will apply and will expire SIX (6) MONTHS from the mailing date of this communication.
- Failure to reply within the set or extended period for reply will, by statute, cause the application to become ABANDONED (35 U.S.C. § 133). Any reply received by the Office later than three months after the mailing date of this communication, even if timely filed, may reduce any earned patent term adjustment. See 37 CFR 1.704(b).

Status

1)☒ Responsive to communication(s) filed on *20 January 2004*.

2a)☐ This action is **FINAL**. 2b)☒ This action is non-final.

3)☐ Since this application is in condition for allowance except for formal matters, prosecution as to the merits is closed in accordance with the practice under *Ex parte Quayle*, 1935 C.D. 11, 453 O.G. 213.

Disposition of Claims

4)☒ Claim(s) *1-14* is/are pending in the application.

4a) Of the above claim(s) _____ is/are withdrawn from consideration.

5)☐ Claim(s) _____ is/are allowed.

6)☒ Claim(s) *1-10* is/are rejected.

7)☐ Claim(s) _____ is/are objected to.

8)☐ Claim(s) _____ are subject to restriction and/or election requirement.

Application Papers

9)☐ The specification is objected to by the Examiner.

10)☒ The drawing(s) filed on *19 October 0199* is/are: a)☒ accepted or b)☐ objected to by the Examiner.

Applicant may not request that any objection to the drawing(s) be held in abeyance. See 37 CFR 1.85(a).

Replacement drawing sheet(s) including the correction is required if the drawing(s) is objected to. See 37 CFR 1.121(d).

11)☐ The oath or declaration is objected to by the Examiner. Note the attached Office Action or form PTO-152.

Priority under 35 U.S.C. § 119

12)☐ Acknowledgment is made of a claim for foreign priority under 35 U.S.C. § 119(a)-(d) or (f).

a)☐ All b)☐ Some * c)☐ None of:

1.☐ Certified copies of the priority documents have been received.

2.☐ Certified copies of the priority documents have been received in Application No. _____.

3.☐ Copies of the certified copies of the priority documents have been received in this National Stage application from the International Bureau (PCT Rule 17.2(a)).

* See the attached detailed Office action for a list of the certified copies not received.

Attachment(s)

1) ☒ Notice of References Cited (PTO-892)
2) ☐ Notice of Draftsperson's Patent Drawing Review (PTO-948)
3) ☐ Information Disclosure Statement(s) (PTO-1449 or PTO/SB/08) Paper No(s)/Mail Date _____.
4) ☐ Interview Summary (PTO-413) Paper No(s)/Mail Date. _____ .
5) ☐ Notice of Informal Patent Application (PTO-152)
6) ☐ Other: _____.

DETAILED ACTION

Note: The examiner for your application in the PTO has changed. However, the Group and/or Art Unit location of your application in the PTO is remained the same, which is Group Art Unit 1639.

Continued Examination Under 37 CFR 1.114

1. A request for continued examination under 37 CFR 1.114, including the fee set forth in 37 CFR 1.17(e), was filed in this application after final rejection. Since this application is eligible for continued examination under 37 CFR 1.114, and the fee set forth in 37 CFR 1.17(e) has been timely paid, the finality of the previous Office action has been withdrawn pursuant to 37 CFR 1.114. Applicant's submission filed on 1/20/04 has been entered.

Status of Claims

2. Claims 1-14 are pending.

3. This application claims priority to a provisional application 60/104,744 filed 10/19/98.

4. Claims 11-14 are withdrawn from further consideration pursuant to 37 CFR 1.142(b) as being drawn to a nonelected inventions, there being no allowable generic or linking claim. Election was made **without** traverse in Paper No. 7 (dated 3/26/01).

5. Claims 1-10 are treated on the merit in this Office Action.

Withdrawn Rejections

6. The previous rejection under 35 USC 112, first paragraph (enablement) has been withdrawn in view of new ground of rejection.

7. The previous rejections under 35 USC 112, second paragraph, have been withdrawn in view of new ground of rejections.

New Rejections

Claim Rejections - 35 USC § 112

8. The following is a quotation of the first paragraph of 35 U.S.C. 112:

> The specification shall contain a written description of the invention, and of the manner and process of making and using it, in such full, clear, concise, and exact terms as to enable any person skilled in the art to which it pertains, or with which it is most nearly connected, to make and use the same and shall set forth the best mode contemplated by the inventor of carrying out his invention.

9. Claims 1-10 are rejected under 35 U.S.C. 112, first paragraph, as containing subject matter which was not described in the specification in such a way as to reasonably convey to one skilled in the relevant art that the inventor(s), at the time the application was filed, had possession of the claimed invention. (This is a written description rejection)

The instant claim 1 recites a method of tag-directed synthesis of a plurality of compounds. The method comprises the steps of a) forming a first group of subsets of nucleic acid tags, b) reacting the chemical reaction sites in each of the subsets formed in (a) with a first selected reagent to form a reagent-specific compound intermediate, c) forming a second group of subsets of the reacted-nucleic acid tags, and d) reacting the compound intermediates in each of the subsets formed in (c) with a second reagent, whereby the nucleic acid tags direct the

synthesis of the compounds. The nucleic acid tags of the first group of subsets comprise a plurality of different first hybridization sequences, a mixture of different second hybridization sequences, and a chemical reaction site. The nucleic acid tags of the second group of subsets comprise a plurality of different second hybridization sequences, and a mixture of different first hybridization sequences.

The specification disclosure does not sufficiently teach the claimed method of tag-directed synthesis using "nucleic acid tag" for synthesizing ***any*** compounds. The specification description is directed to a method of synthesizing DNA library specifically base-specific duplex formation using "nucleic acid tag" (pg. 12, lines 5-8; pg. 14, line 9 to pg. 15, line 22). Thus the specification does not teach the claimed method of tag-directed synthesis using "nucleic acid tag" for synthesizing ***any*** compounds such as peptides, oligosaccharides, or antibody.

The specification disclosure does not sufficiently teach the claimed method of tag-directed synthesis wherein the nucleic acid tags of the first group of subsets comprise a plurality of different first hybridization sequences, a mixture of different second hybridization sequences, and a chemical reaction site, and the nucleic acid tags of the second group of subsets comprise a plurality of different second hybridization sequences, and a mixture of different first hybridization sequences. The specification description discloses ***one*** group of subsets of nucleic acid tags (pg. 11, lines 24-29; pg. 12, line 22 to pg. 13, line 21; pg. 14, lines 29-32). The specification specifically discloses that "***each tag*** has ***a first segment*** having a selected one of a plurality of different first hybridization sequences, a mixture of different second hybridization sequences, and a chemical reaction site; ***and a second segment*** having a selected one of a plurality of different second hybridization sequences and a mixture of different first

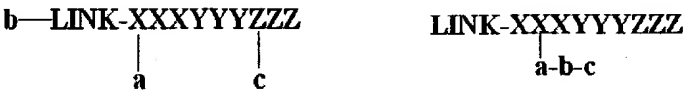
hybridization sequences" (pg. 11, lines 24-29). Therefore, the specification does not teach the claimed method of tag-directed synthesis wherein the nucleic acid tags of the first group of subsets comprise a plurality of different first hybridization sequences, a mixture of different second hybridization sequences, and a chemical reaction site, and the nucleic acid tags of the second group of subsets comprise a plurality of different second hybridization sequences, and a mixture of different first hybridization sequences.

The specification disclosure does not sufficiently teach the claimed method of tag-directed synthesis wherein the method comprises the steps of forming a second group of subsets of the reacted-nucleic acid tags, and reacting the compound intermediates in each of the second group of subsets formed with a second reagent. The specification description is directed to forming ***one*** group of subsets of nucleic acid tags (pg. 11, lines 24-29; pg. 12, line 22 to pg. 13, line 21; pg. 14, lines 29-32). The specification is silent on the method steps of forming a second group of subsets of the reacted-nucleic acid tags, and reacting the compound intermediates in each of the second group of subsets formed with a second reagent.

The specification disclosure does not sufficiently teach the claimed method of tag-directed synthesis wherein the method step of attaching a reagent to the "nucleic acid " tag encompasses ***any*** type of reaction mechanisms. For example, the "nucleic acid " tag comprises LINK-XXXYYYZZZ, wherein LINK denotes either a solid support with a functional group (chemical reaction site) or a chemical linker with a functional group (chemical reaction site), and 'X', 'Y', and 'Z' denotes nucleic acid that comprises several different "chemical reaction site" such as nucleotide base, sugar, or the backbone. Thus there is several different mode of

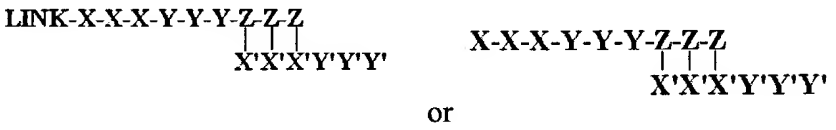
attaching the reagent to the "nucleic acid " tag. Some examples of the reaction mechanisms for attaching reagents are as follows: (reagents are denoted as 'a', 'b', and 'c')

- d-c-b-a- LINK-XXXYYYZZZ;
- LINK-XXXYYYZZZ-a-b-c-d;



- Branching such as: or
- Extending the length of the 'nucleic acid' tag such as:

LINK-XXXYYYZZZX'X'Y'Y'Y'; or by hybridization such as



In each of these reaction mechanisms the chemistry required for attachments of the reagent(s) is constrained by the type of protecting group use and the method of deprotection. The specification description is directed to a method of synthesizing DNA library specifically base-specific duplex formation using "nucleic acid tag" (pg. 12, lines 5-8; pg. 14, line 9 to pg. 15, line 22). Thus the specification does not teach the claimed method of tag-directed synthesis wherein the method step of attaching a reagent to the "nucleic acid " tag encompasses ***any*** type of reaction mechanisms other than hybridization to form base-specific duplex.

Vas-Cath Inc. v. Mahurkar, 19 USPQ2d 1111, makes clear that "applicant must convey with reasonable clarity to those skilled in the art that, as of the filing date sought, he or she was in possession *of the invention.* The invention is, for purposes of the 'written description' inquiry, *whatever is now claimed.*" (See Vas-Cath at page 1117.) The specification does not "clearly

allow persons of ordinary skill in the art to recognize that [he or she] invented what is claimed." (See Vas-Cath at page 1116.).

With the exception of the method of synthesizing DNA library specifically base-specific duplex formation using "nucleic acid tag" disclosed by the specification, the skilled artisan cannot envision the claimed method of tag-directed synthesis wherein the method would a) synthesize ***any*** compounds, b) encompasses ***any*** type of reaction mechanisms c) comprise the steps of forming a second group of subsets of the reacted-nucleic acid tags, and reacting the compound intermediates in each of the second group of subsets formed with a second reagent, and d) forming two groups of nucleic acid tags. Adequate written description requires more than a mere statement that it is part of the invention and/or reference to a potential method for isolating it. See Fiers v. Revel, 25 USPQ2d 1601, 1606 (CAFC 1993) and Amgen Inc. V. Chugai Pharmaceutical Co. Ltd., 18 USPQ2d 1016. In Fiddes v. Baird, 30 USPQ2d 1481, 1483, claims directed to mammalian FGF's were found unpatentable due to lack of written description for the broad class. The specification provided only the bovine sequence.

Finally, University of California v. Eli Lilly and Co., 43 USPQ2d 1398, 1404, 1405 held that:

...To fulfill the written description requirement, a patent specification must describe an invention and do so in sufficient detail that one skilled in the art can clearly conclude that "the inventor invented the claimed invention." *Lockwood v. American Airlines, Inc.*, 107 F.3d 1565, 1572, 41 USPQ2d 1961, 1966 (1997); In *re Gosteli*, 872 F.2d 1008, 1012, 10 USPQ2d 1614, 1618 (Fed. Cir. 1989) (" [T]he description must clearly allow persons of ordinary skill in the art to recognize that [the inventor] invented what is claimed."). Thus, an applicant complies with the written description requirement "by describing the invention, with all its claimed limitations, not that which makes it obvious," and by using "such descriptive means as words, structures, figures, diagrams, formulas, etc., that set forth the claimed invention." *Lockwood*, 107 F.3d at 1572, 41 USPQ2d at 1966.

In the present instance, the specification supports the method of synthesizing DNA library specifically base-specific duplex formation using "nucleic acid tag". The specification does not teach the claimed method of tag-directed synthesis wherein the method would a) synthesize ***any*** compounds, b) encompasses ***any*** type of reaction mechanisms c) comprise the steps of forming a second group of subsets of the reacted-nucleic acid tags, and reacting the compound intermediates in each of the second group of subsets formed with a second reagent, and d) forming two groups of nucleic acid tags. Therefore, only the method of synthesizing DNA library specifically base-specific duplex formation using "nucleic acid tag", but not the full breadth of the claimed method of tag-directed synthesis wherein the method would a) synthesize ***any*** compounds, b) encompasses ***any*** type of reaction mechanisms c) comprise the steps of forming a second group of subsets of the reacted-nucleic acid tags, and reacting the compound intermediates in each of the second group of subsets formed with a second reagent, and d) forming two groups of nucleic acid tags meet the written description provision of 35 U.S.C 112, first paragraph.

10. The following is a quotation of the second paragraph of 35 U.S.C. 112:

> The specification shall conclude with one or more claims particularly pointing out and distinctly claiming the subject matter which the applicant regards as his invention.

11. Claims 1-10 are rejected under 35 U.S.C. 112, second paragraph, as being indefinite for failing to particularly point out and distinctly claim the subject matter which applicant regards as the invention.

a. Claim 1 recites "a chemical reaction site in each of the subsets of the nucleic acid tags" is indefinite because it is unclear to the means of determining the "location" of the reaction site on the "nucleic acid tag' (i.e. where on the tag is the reagent being attached).

b. The recitation of "the reacted-nucleic acid tags" of claim 1 is vague and indefinite because it is unclear if it is referring to the "reagent-specific compound intermediate formed in step (b) or the 'product' of the claimed method. Furthermore, step (d) refers to "the reacted-nucleic acid tags" as the 'compound intermediates'.

c. Claim 1 is rejected under 35 U.S.C. 112, second paragraph, as being incomplete for omitting essential steps, such omission amounting to a gap between the steps. See MPEP § 2172.01. The omitted steps are:

1) The step between step (b) forming the 'reagent-specific compound intermediate' and step (c) forming the second group of tags because it is unclear as to the correlation between these to step (i.e. what happen to the 'reagent-specific compound intermediate').

2) The step between step (b) forming the 'reagent-specific compound intermediate' and step (d) forming the compounds of the claimed method because it is unclear as to the "role" in which the 'reagent-specific compound intermediate' play in forming the compounds of the claimed method since the combination of method steps (c) and (d) would produce the "compounds" where as the combination of method steps (a) and (b) would produce the "compound intermediate".

d. The term 'reagents' of claim 2, line 4 is indefinite because it is unclear whether this reagent refers to the "solid phase reagent" of the surface bound oligonucleotide or the reagents of claim 1 in step (b) and/or (c).

e. Claim 5 recites the limitation "synthetic step" in lines 1-2. There is insufficient antecedent basis for this limitation in the claim 1.

f. The recitation of "a selected subunit" of claim 3 and "a selected chemical substituent" of claim 4 are vague and indefinite because it is unclear if the "subunit" and "chemical substituent" or are they referring to the reagents of claim 1 (i.e. first reagent and second reagent). Furthermore, it is unclear if there is a relationship (i.e. structural and/or functional) among these 'compounds' (i.e. "subunit", "chemical substituent", "first reagent" and "second reagent")

Claim Rejections - 35 USC § 102

12. The following is a quotation of the appropriate paragraphs of 35 U.S.C. 102 that form the basis for the rejections under this section made in this Office action:

> A person shall be entitled to a patent unless –
>
> (b) the invention was patented or described in a printed publication in this or a foreign country or in public use or on sale in this country, more than one year prior to the date of application for patent in the United States.
>
> (e) the invention was described in a patent granted on an application for patent by another filed in the United States before the invention thereof by the applicant for patent, or on an international application by another who has fulfilled the requirements of paragraphs (1), (2), and (4) of section 371(c) of this title before the invention thereof by the applicant for patent.

The changes made to 35 U.S.C. 102(e) by the American Inventors Protection Act of 1999 (AIPA) and the Intellectual Property and High Technology Technical Amendments Act of 2002 do not apply when the reference is a U.S. patent resulting directly or indirectly from an

international application filed before November 29, 2000. Therefore, the prior art date of the reference is determined under 35 U.S.C. 102(e) prior to the amendment by the AIPA (pre-AIPA 35 U.S.C. 102(e)).

13. Claims 1, and 3-6 are rejected under 35 U.S.C. 102(b) as being anticipated by Brenner et al. (*Proc. Natl. Acad. Sci.*, **1992**, 89(12):5381-5383).

Brenner et al. disclose the method of two alternating parallel combinatorial syntheses wherein the genetic tag is chemically link to the chemical structure being synthesized (pg. 5381, right col., lines 47-54). In general, the method comprises the addition of monomeric chemical unit to a polymeric structure that is followed by the addition of an oligonucleotide sequence, which is defined a "encoding" that chemical unit (tag-directed synthesis). The library of compounds is built up by the repetition of this process after pooling and division. The method comprises the steps of 1) synthesizing the first PCR oligonucleotide sequence on one end of the solid support, 2) dividing the PCR oligonucleotide bound to a support into two aliquots for parallel synthesis, 3) adding to each aliquots an amino acid (first reagent/subunit) and an oligonucleotide coding sequence (chemical substituent) for the amino acid wherein the amino acid is attached to the other end of the solid support and the oligonucleotide coding sequence is attached to the PCR oligonucleotide, which would extend the PCR oligonucleotide sequence (refers to step (b) of claim 1 and claims 3-4), 4) the two aliquots are pooled and again split for parallel synthesis in which step (3) is repeated to form the desired elongated product (reagent-specific compound intermediate) (refers to claim 5), and 5) attaching the second PCR oligonucleotide to the elongated product to produce a library of compounds (pg. 5382, right col.,

lines 34-66). The compound comprises a genetic tag, a solid support, and peptide sequences. The elongated product (first group) comprises a plurality of oligonucleotide coding sequences (a mixture of different second hybridization sequences), the first PCR oligonucleotide sequence (first hybridization sequence), and a solid support (chemical reactive site) wherein the amino acid is attached. Additionally, Brenner et al. disclose that DNA strands with the appropriate polarity can then be used to enrich for a subset of the library by hybridization with matching tags, and the process can then be repeated on this subset (refers to steps (c) and (d) of claim 1). Thus the method of Brenner et al. anticipates the presently claimed method.

14. Claims 1-10 are rejected under 35 U.S.C. 102(b) as being anticipated by Brenner (US Patent 5,604,097).

Brenner discloses a method of sorting polynucleotides from a mixture of polynucleotides by hybridizing the polynucleotides to their complements of oligonucleotide tags (tag-directed synthesis) (col. 3, lines 33-58). Each of the polynucleotides (first group of subsets of nucleic acid tags) of a mixture of polynucleotides to be sorted comprises an oligonucleotide tag in the repertoire such that identical polynucleotides have the same tag and different polynucleotides have different tags (col. 3, lines 48-51). The oligonucleotide tags (second group of subsets of nucleic acid tags) comprise a repertoire of complements with distinct sequences wherein the size of the repertoire depends on the number of subunits and length of subunits employed (col. 3, lines 44-48). The method of synthesizing the polynucleotide of a mixture of polynucleotides comprises the steps of 1) an oligonucleotide segment (first hybridization sequence) is synthesized initially so that in double stranded form a restriction endonucleoase site (adjacent

spacer sequence) is provided for cleaving the library compound after sorting, and 2) adding alternating subunits to form the oligonucleotide tag (a mixture of different second hybridization sequences) and their corresponding compound monomers (reagent) via a "split and mix" technique (col. 12, lines 5-14) (refers to steps (a) and (b) of claim 1). The oligonucleotide tags are synthesized on a solid phase support by subunit-wise synthesis via "split and mix" technique (col. 3, lines 33-58; col. 8, lines 16-37) (refers to step (c) of claim 1). The oligonucleotide tag and the mixture of polynucleotide are reacted to form perfectly formed duplexes by hybridization (col. 16, lines 1-19) (refers to step (d) of claim 1 and claim 2). Additionally, the duplexes are identified by way of polynucleotide sequencing (col. 19, line 52 to col. 20, line 14). Therefore the method of Brenner anticipates the presently claimed method.

15. Claims 1, and 7-10 are rejected under 35 U.S.C. 102(e) as being anticipated by Brenner (US Patent 5,962,228).

Brenner discloses a method of synthesizing polynucleotides from a mixture of polynucleotides by hybridizing the polynucleotides to their complements of oligonucleotide tags (tag-directed synthesis) (col. 3, lines 26-46). Each of the polynucleotides (first group of subsets of nucleic acid tags) of a mixture of polynucleotides to be sorted comprises an oligonucleotide tag in the repertoire such that identical polynucleotides have the same tag and different polynucleotides have different tags (col. 4, lines 45-59) (refers to steps (a) and (b) of claim 1). The oligonucleotide tags (second group of subsets of nucleic acid tags) comprise a repertoire of complements with distinct sequences wherein the size of the repertoire depends on the number of subunits and length of subunits employed (col. 4, lines 14-44). The oligonucleotide tags are

synthesized on a solid phase support by subunit-wise synthesis via "split and mix" technique (col. 4, lines 45-59; col. 13, lines 7-13) (refers to step (c) of claim 1). Additionally, the duplexes are identified by way of polynucleotide sequencing (col. 17, line 27 to col. 20, line 59). Therefore the method of Brenner anticipates the presently claimed method.

Conclusion

Any inquiry concerning this communication or earlier communications from the examiner should be directed to MY-CHAU T TRAN whose telephone number is 571-272-0810. The examiner can normally be reached on Mon.: 8:00-2:30; Tues.-Thurs.: 7:30-5:00; Fri.: 8:00-3:30.

If attempts to reach the examiner by telephone are unsuccessful, the examiner's supervisor, ANDREW WANG can be reached on 571-272-0811. The fax phone number for the organization where this application or proceeding is assigned is 703-872-9306.

Information regarding the status of an application may be obtained from the Patent Application Information Retrieval (PAIR) system. Status information for published applications may be obtained from either Private PAIR or Public PAIR. Status information for unpublished applications is available through Private PAIR only. For more information about the PAIR system, see http://pair-direct.uspto.gov. Should you have questions on access to the Private PAIR system, contact the Electronic Business Center (EBC) at 866-217-9197 (toll-free).

mct

May 14, 2004

PADMASHRI PONNALURI
PRIMARY EXAMINER

Proc. Natl. Acad. Sci. USA
Vol. 89, pp. 5381–5383, June 1992
Chemistry

Encoded combinatorial chemistry

(chemical repertoire/encoded libraries/commaless code)

SYDNEY BRENNER AND RICHARD A. LERNER

Departments of Chemistry and Molecular Biology, The Scripps Research Institute, 10666 North Torrey Pines, La Jolla, CA 92037

Contributed by Sydney Brenner, March 3, 1992

ABSTRACT **The diversity of chemical synthesis and the power of genetics are linked to provide a powerful, versatile method for drug screening. A process of alternating parallel combinatorial synthesis is used to encode individual members of a large library of chemicals with unique nucleotide sequences. After the chemical entity is bound to a target, the genetic tag can be amplified by replication and utilized for enrichment of the bound molecules by serial hybridization to a subset of the library. The nature of the chemical structure bound to the receptor is decoded by sequencing the nucleotide tag.**

There is an increasing need to find new molecules that can effectively modulate a wide range of biological processes, for applications in medicine and agriculture. A standard way to search for novel chemicals is to screen collections of natural materials, such as fermentation broths, plant extracts, or libraries of synthesized molecules. Assays can range in complexity from simple binding reactions to elaborate physiological preparations. The screens often only provide leads, which then require further improvement either by empirical methods or by chemical design. The process is time-consuming and costly but is unlikely to be replaced totally by rational methods even when they are based on detailed knowledge of the three-dimensional structure of the target molecules. Thus, what we might call "irrational drug design"—the process of selecting the correct molecules from large ensembles or repertoires—requires continual improvement both in the generation of repertoires and in the methods of selection.

Recently there have been several developments in using peptides or nucleotides to provide libraries of compounds for discovery of leads. The methods were originally developed to speed up the determination of epitopes recognized by monoclonal antibodies. For example, the standard serial process of stepwise search of synthetic peptides has been replaced by a variety of highly sophisticated methods in which large arrays of peptides are synthesized in parallel and screened with acceptor molecules labeled with fluorescent or other reporter groups (1, 2). The sequence of any effective peptide can be decoded from its address in the array. In another approach, combinatorial libraries of peptides are synthesized on resin beads such that each resin bead contains about 20 pmol of the same peptide (3). The beads are exposed to labeled acceptor molecules. Those with bound acceptor are identified by visual inspection and physically removed, and the peptide is sequenced directly. In principle, this method could be used with other chemical entities, provided one has a sensitive method for sequence determination.

A different method of solving the problem of identification in a combinatorial peptide library is used by Houghten *et al.* (4). For hexapeptides of the 20 natural amino acids, separate libraries are synthesized, each with the first two amino acids fixed and the remaining four positions occupied by all possible combinations. An assay, based on competition for binding or some other activity, is then used to find the library with an active peptide. On the basis of this result, 20 new libraries are synthesized and assayed to determine the effective amino acid in the third position. The process is reiterated in this fashion until the active hexapeptide is defined. This is analogous to the method used in searching a dictionary: the peptide is decoded by using a series of sieves, and this makes the search logarithmic. A powerful biological method has recently been described in which the library of peptides is presented on the surface of a bacteriophage such that each phage displays a particular peptide and contains within its genome the corresponding DNA sequence (5, 6). The library is prepared by synthesizing a repertoire of random oligonucleotides to generate all combinations, followed by their insertion into a phage vector. Each of the sequences is cloned in one phage and the relevant peptide can be selected by finding those that bind to the particular target. The phages recovered in this way can be amplified and the selection repeated. The sequence of the peptide is decoded by sequencing the DNA. Another "genetic" method has been applied by Tuerk and Gold (7) and Ellington and Szostak (8), using libraries of synthetic oligonucleotides that themselves are selected for binding to an acceptor and then amplified by the polymerase chain reaction (PCR). In this case, however, the repertoire is limited to nucleotides or nucleotide analogues that preserve specific Watson–Crick pairing and can be copied by a polymerase.

The main advantages of the genetic methods reside in the capacity for cloning and amplification of DNA sequences, which allows enrichment by serial selection and provides a facile method for decoding the structure of active molecules. However, the genetic repertoires are restricted to nucleotides and peptides composed of natural amino acids, whereas a more extensive chemical repertoire is required to populate the entire universe of binding sites. In contrast, chemical methods can provide limitless repertoires, but they lack the capacity for serial enrichment and there are difficulties in discovering the structures of selected active molecules. We have now devised a way of combining the virtues of both methods through the construction of *encoded combinatorial chemical libraries*, in which each chemical sequence is labeled by an appended "genetic" tag, itself constructed by chemical synthesis. In effect, we implement a "retrogenetic" way of specifying each chemical structure.

In outline, we perform two alternating parallel combinatorial syntheses so that the genetic tag is chemically linked to the chemical structure being synthesized. In each case, addition of a monomeric chemical unit to a polymeric structure is followed by addition of an oligonucleotide sequence which is defined as "encoding" that chemical unit. The library is built up by the repetition of this process after pooling and division. Active molecules are selected by binding to a receptor, and amplified copies of their retrogenetic tags are obtained by the PCR. DNA strands with the appropriate polarity can then be used to enrich for a subset of the

The publication costs of this article were defrayed in part by page charge payment. This article must therefore be hereby marked "*advertisement*" in accordance with 18 U.S.C. §1734 solely to indicate this fact.

library by hybridization with the matching tags, and the process can then be repeated on this subset. Thus serial enrichment is achieved by a process of purification, exploiting linkage to a nucleotide sequence that can be amplified. Finally, the structures of the chemical entities are decoded by cloning and sequencing the products of PCR.

Design of the Code and the Genetic Tag

It is essential to choose a coding representation in such a way that no significant part of the sequence can occur by chance in some other unrelated combination. Suppose we allocate a triplet to each of the chemical units used. Then, because the method allows us to cover all combinations and permutations of an alphabet of chemical units, unless we are careful, we could find that two different combinations have closely related sequences which differ only by a frame shift and which could not be easily distinguished by hybridization. This, potentially the greatest source of errors, can be eliminated by choosing a commaless code (9). The particular commaless triplet code that we have chosen allows 20 unique representations, as shown in Table 1.

The sequences for the PCR primers must be chosen so that they do not occur within any coding segment and so that they can be readily removed from the final PCR product because we do not want them to dominate the selective hybridization. This can be achieved by building in sites for restriction enzymes with the appropriate polarity of cutting. One of the restriction enzymes should cut at a site that permits the incorporation of a biotinylated nucleotide, such as biotinyl-dUTP, into the strand complementary to the coding strand.

All of the above conditions have been met in the following design:

5′-AGCTACTTCCCAAGG[coding sequence]GGGCCCTATTCTTAG-3′
3′-TCGATGAAGGGTTCC[anticoding strand]CCCGGGATAAGAATC-5′
Sty I *Apa* I

After cleavage with both restriction enzymes we have

5′-AGCTACTTCC CAAGG[coding sequence]GGGCC CTATTCTTAG-3′
3′-TCGATGAAGGGTTC C[anticoding strand]C CCGGGATAAGAATC-5′

The internal fragment can be cloned in an appropriate vector to sequence the individuals. The terminal overhang of the *Sty* I site can be filled in with dCTP and biotinyl-dUTP (BTP) which, because an asymmetric site was chosen, will append the biotinylated nucleotides to only one of the cleavage products.

5′-AGCTACTTCCC CAAGG[coding sequence]GGGCC CTATTCTTAG-3′
3′-TCGATGAAGGGTTC BBCC[anticoding strand]C CCGGGATAAGAATC-5′

The biotinylated fragment can be bound to avidin and, after denaturation, provides the strand suitable for hybridization and selection of the appropriate coding strands:

Avidin-BBCC[anticoding strand]C

The two PCR primers are the two sequences 5′-AGCTACTTCCCAAGG (*Sty* I primer) and 5′-CTAAGAATAGGGCCC (*Apa* I primer). Adding a biotin to the 5′ end of the *Apa* I primer would allow the isolation of the whole strand containing the anticoding sequence.

We should have at least 15 nucleotides in the coding region for effective hybridization. Thus, in a library of degree $d \geq 5$, that is, composed of five or more successive chemical units, we could code each unit by a triplet. That would allow an alphabet (A) of up to 20 different units, each corresponding to one of the triplets defined above. The complexity of the combinatorial library is A^d. Libraries with a smaller degree, say $d = 3$, should be coded by sextuplets, which, in the simplest case, could be a repeated triplet (this size is chosen because any combination of triplets still obeys the commaless condition). In the same way, the size of the alphabet can be extended by using combinations of triplets to code for the chemical units.

A Formal Example

As an illustration we discuss how a library of degree $d = 3$ is made with an alphabet of two amino acids, glycine and methionine. In this case, we use sextuplets to give us a reasonable length of coding sequence. To make the sequences as different as possible we code each amino acid by a combination of two different triplets as follows:

Gly = CACATG, Met = ACGGTA.

Step 1. We begin with some appropriate linker, LINK, attached to some solid-state surface and synthesize the first PCR oligonucleotide sequence on one end, in the usual 3′-to-5′ direction, to give

GGGCCCTATTCTTAG–LINK

Step 2. This product is divided into two aliquots for parallel synthesis. In each synthesis, one amino acid is added to LINK and the oligonucleotide sequence is extended by the corresponding code to give the following products:

CACATGGGGCCCTATTCTTAG–LINK–Gly
ACGGTAGGGCCCTATTCTTAG–LINK–Met

Step 3. The elongated products are pooled and again split into two parts for parallel synthesis, yielding

CACATGCACATGGGGCCCTATTCTTAG–LINK–Gly–Gly
CACATGACGGTAGGGCCCTATTCTTAG–LINK–Met–Gly
ACGGTACACATGGGGCCCTATTCTTAG–LINK–Gly–Met
ACGGTAACGGTAGGGCCCTATTCTTAG–LINK–Met–Met

Steps 4 and 5. Once more the products are pooled and divided into two aliquots for parallel synthesis. This results in an ensemble of eight tripeptide sequences, each encoded by a unique sequence of 18 nucleotides. The second PCR oligonucleotide is added to the ensemble of products to give

Table 1. Commaless code used in this study

ttt	tct	tat	tgt
TTC	tcc	tac	tgc
TTA	tca	taa	tga
TTG	tcg	tag	tgg
ctt	cct	cat	cgt
CTC	ccc	cac	cgc
CTA	cca	caa	cga
CTG	ccg	cag	cgg
att	act	aat	agt
ATC	ACC	aac	agc
ATA	ACA	aaa	aga
ATG	ACG	aag	agg
gtt	gct	gat	ggt
GTC	GCC	gac	ggc
GTA	GCA	GAA	gga
GTG	GCG	GAG	ggg

"Sense triplets" are XYZ; nonsense triplets are xyz.

```
AGCTACTTCCCAAGGCACATGCACATGCACATGGGGCCCTATTCTTAG-LINK-Gly-Gly-Gly
AGCTACTTCCCAAGGCACATGCACATGACGGTAGGGCCCTATTCTTAG-LINK-Met-Gly-Gly
AGCTACTTCCCAAGGCACATGACGGTACACATGGGGCCCTATTCTTAG-LINK-Gly-Met-Gly
AGCTACTTCCCAAGGCACATGACGGTAACGGTAGGGCCCTATTCTTAG-LINK-Met-Met-Gly
AGCTACTTCCCAAGGACGGTACACATGCACATGGGGCCCTATTCTTAG-LINK-Gly-Gly-Met
AGCTACTTCCCAAGGACGGTACACATGACGGTAGGGCCCTATTCTTAG-LINK-Met-Gly-Met
AGCTACTTCCCAAGGACGGTAACGGTACACATGGGGCCCTATTCTTAG-LINK-Gly-Met-Met
AGCTACTTCCCAAGGACGGTAACGGTAACGGTAGGGCCCTATTCTTAG-LINK-Met-Met-Met
```

Implementation

Although natural amino acids are used in the example discussed above, the system is not limited to these, nor, for that matter, to peptides. The chemistry required for making encoded libraries is constrained only by the compatibility of the two alternating syntheses. Partly this involves the choice of the protecting groups, and the methods used to deprotect one chain while the other remains blocked. And, of course, each product needs to survive through the synthesis of the other. One can imagine many different ways of joining the chemical entities together, and one could even use mixed syntheses, provided that the rules of mutual compatibility are obeyed.

We have recently, in principle, solved the synthetic procedures for peptides (K. Janda, S. Ramcharitar, S.B., and R.A.L., unpublished results). Even within this field there is a choice of alphabets that extends well beyond the 20 natural α-amino acids. The only requirement is that we be able to make an amide bond. Thus, the amino and carboxylic groups can be located on a wide variety of compounds so that we can make libraries with many different backbone structures. We can also combine different backbones, if we define alphabets where, for example, both the number of carbon atoms and their configurations in the backbone are varied. New amino acids can be easily invented with unusual heterocyclic rings, such as thiazole-alanine or purine-alanine. These rings are components of natural effector molecules and often provide core chemical functions for important drugs. Libraries made with such alphabets will allow us to explore the combinatorial association of known effector chemical functions.

It is also useful to consider how large the combinatorial library should be. The PCR provides a very sensitive detection method, allowing even a few molecules to be seen. However, we need to have some reasonable concentration of each of the species present to cross the binding threshold of the acceptor molecule being assayed. If, for example, we set this as 1 μM and want 1 ml of the library, then we need to make at least 1 nmol of each of the species. Libraries with complexities of up to 10^4, giving us a total amount of 10 μmol of product, would seem reasonable. Because of this reciprocal relationship, more complex libraries could be made if the binding threshold is lowered.

Discussion

Traditional chemical synthesis proceeds by careful design, sequentially linking atoms or groups of atoms to a growing core structure. The process has the advantage that the product of each step can be analyzed, thereby allowing continuous evaluation of the effectiveness of a given strategy. Indeed, the analyzed results of these individual steps ultimately become the corpus of synthetic organic chemistry. A major technical revolution occurred with the advent of solid-state methods for the synthesis of polymeric molecules (10). Here, since a limited number of suitably protected oligomeric units are added via a common covalent bond, the results of the individual transformations can be predicted, and, to first approximation, it is necessary to analyze only the final product. In addition, the relationship of the monomeric units to each other and the extent of conformational space that is occupied can be estimated. Our method permits the study of the efficacy of combinatorial associations of diverse chemical units without the necessity of either synthesizing them one at a time or knowing their interactions in advance. It also allows easy identification of the most effective molecules through a common method of nucleic acid sequencing. Once the chemical polymers are decoded, more precise questions about critical interactions and conformations can be asked by reversion to classical chemical methods. Further, we expect that many receptors will interact with sets of related but not identical chemical entities such that major clues as to critical interactions can be deduced from the shared features of the sets.

Our method also provides a method of amplification, again by exploiting a common procedure of nucleic acid hybridization. In any screening procedure where large libraries of compounds or effector molecules are being studied, the absolute number of different nonspecific interactions may be large, but the specific ligand or effector is represented many more times than any individual background molecule. In such a situation the signal-to-noise ratio rapidly increases after repeated cycles of amplification and selection, and the specific molecule becomes highly enriched after only a few iterations. For both identification and selection, our method exploits the power of genetic systems. By coupling genetics and the versatility of organic chemical synthesis we have extended the range of analysis to chemicals that are not themselves part of biological systems.

We thank Kim Janda, Bernie Gilula, and Jerry Joyce for helpful comments on the manuscript.

1. Fodor, S. P. A., Read, J. L., Pirrung, M. C., Stryer, L., Tsai, Lu, A. & Solas, D. (1991) *Science* **251**, 767–773.
2. Geysen, H. M., Meloen, R. H. & Barteling, S. J. (1984) *Proc. Natl. Acad. Sci. USA* **81**, 3998–4002.
3. Lam, K. S., Salmon, S. E., Hersh, E. M., Hruby, V. J., Kazmierski, W. M. & Knapp, R. J. (1991) *Nature (London)* **354**, 82–84.
4. Houghten, R. A., Pinilla, C., Blondelle, S. E., Appel, J. R., Dooley, C. T. & Cuervo, J. H. (1991) *Nature (London)* **354**, 84–86.
5. Scott, J. K. & Smith, G. P. (1990) *Science* **249**, 386–390.
6. Cwirla, S. E., Peters, E. A., Barrett, R. W. & Dower, W. J. (1990) *Proc. Natl. Acad. Sci. USA* **87**, 6378–6382.
7. Tuerk, C. & Gold, L. (1990) *Science* **24**, 505–510.
8. Ellington, A. D. & Szostak, J. W. (1990) *Nature (London)* **346**, 818–822.
9. Crick, F. H. C., Griffith, J. S. & Orgel, L. E. (1957) *Proc. Natl. Acad. Sci. USA* **43**, 416–421.
10. Merrifield, B. (1984) *Les Prix Nobel* (Almqvist & Wiksell International, Stockholm), pp. 127–153.

Notice of References Cited	Application/Control No. 09/421,422	Applicant(s)/Patent Under Reexamination HARBURY ET AL.	
	Examiner MY-CHAU T TRAN	Art Unit 1639	Page 1 of 1

U.S. PATENT DOCUMENTS

*		Document Number Country Code-Number-Kind Code	Date MM-YYYY	Name	Classification
	A	US-5,604,097	02-1997	Brenner, Sydney	435/6
	B	US-5,962,228	10-1999	Brenner, Sydney	435/6
	C	US-			
	D	US-			
	E	US-			
	F	US-			
	G	US-			
	H	US-			
	I	US-			
	J	US-			
	K	US-			
	L	US-			
	M	US-			

FOREIGN PATENT DOCUMENTS

*		Document Number Country Code-Number-Kind Code	Date MM-YYYY	Country	Name	Classification
	N					
	O					
	P					
	Q					
	R					
	S					
	T					

NON-PATENT DOCUMENTS

*		Include as applicable: Author, Title Date, Publisher, Edition or Volume, Pertinent Pages)
*	U	Brenner et al., "Encoded combinatorial chemistry", 1992, Proc. Natl. Acad. Sci., 89(15):5381-5383.
	V	
	W	
	X	

*A copy of this reference is not being furnished with this Office action. (See MPEP § 707.05(a).)
Dates in MM-YYYY format are publication dates. Classifications may be US or foreign.